Comparison of Cloud Approaches

Overview and Perspectives from STAR

Levente HAJDU
Lidia DIDENKO
Jérôme LAURET
Doug OLSON

STAR

Open Science Grid

U.S. DEPARTMENT OF ENERGY

Office of Science
U.S. DEPARTMENT OF ENERGY

International Symposium on Grid Computing
ISGC

Outline

- Intro
- Concepts:
  - Contextualization & Base Images
  - VM Image Management
  - Efficiency
- Models we have run in production (pros and cons):
  - Non-Virtualized VDT/OSG Model
  - Amazon EC2 with Nimbus interface - Totally Virtualized grid site
  - Clemson Model CI#1 - Virtualized worker nodes, with batch worker daemon inside
  - VM Model G#1- Virtualized VM started by external batch worker
- What would be the ideal model ?

※Naturally all sites upgrade and improve their operating models over time. What we are presenting here is a snapshot in time of what we have observed from Clouds STAR has produced data on.

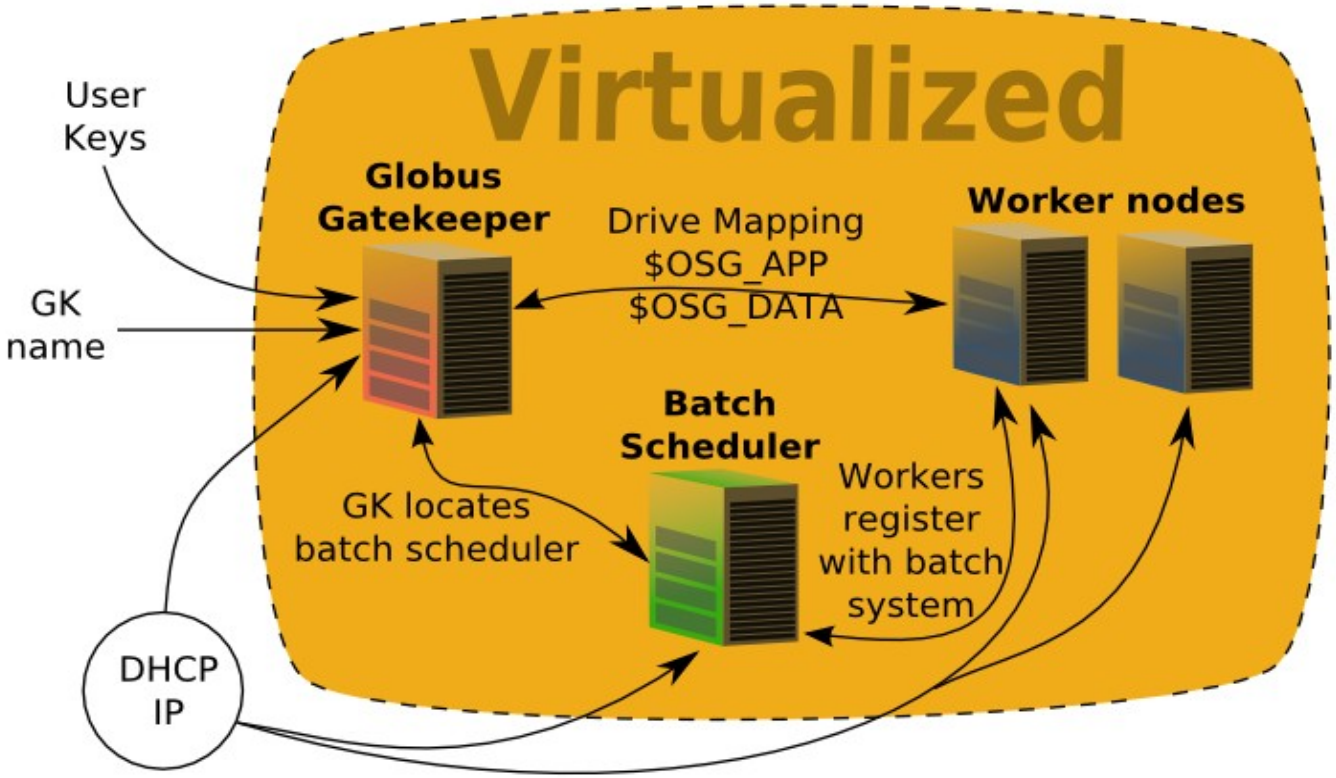
Introduction

- **Cloud Computing** is an emerging trend
  - Multiple providers: from Amazon EC2, Magellan (DOE), Azure Cloud (NSF), SGI Cyclone, ...
  - Multiple software stacks and approaches: Nimbus, Eucalyptus, Cloudera, ...
- Is there a way to merge Cloud and Grids?
  - Or can Grid gain from Cloud "philosophy"?
- STAR's work
  - STAR has run physics jobs at different facilities for the purpose of Evaluating different approaches and designs
  - Presentation of pro and con study in a scientific computing context (some approach will be easier for end users, some easier for administrators)
- * Why?
  - Virtualization providing an easy way toward environment and software provisioning, interest in "a" solution is high.
    - Guarantees reproducibility of results

Contextualization & Base Images

Contextualization is initialization that is required at or after VM image boot time, before any jobs can be submitted.

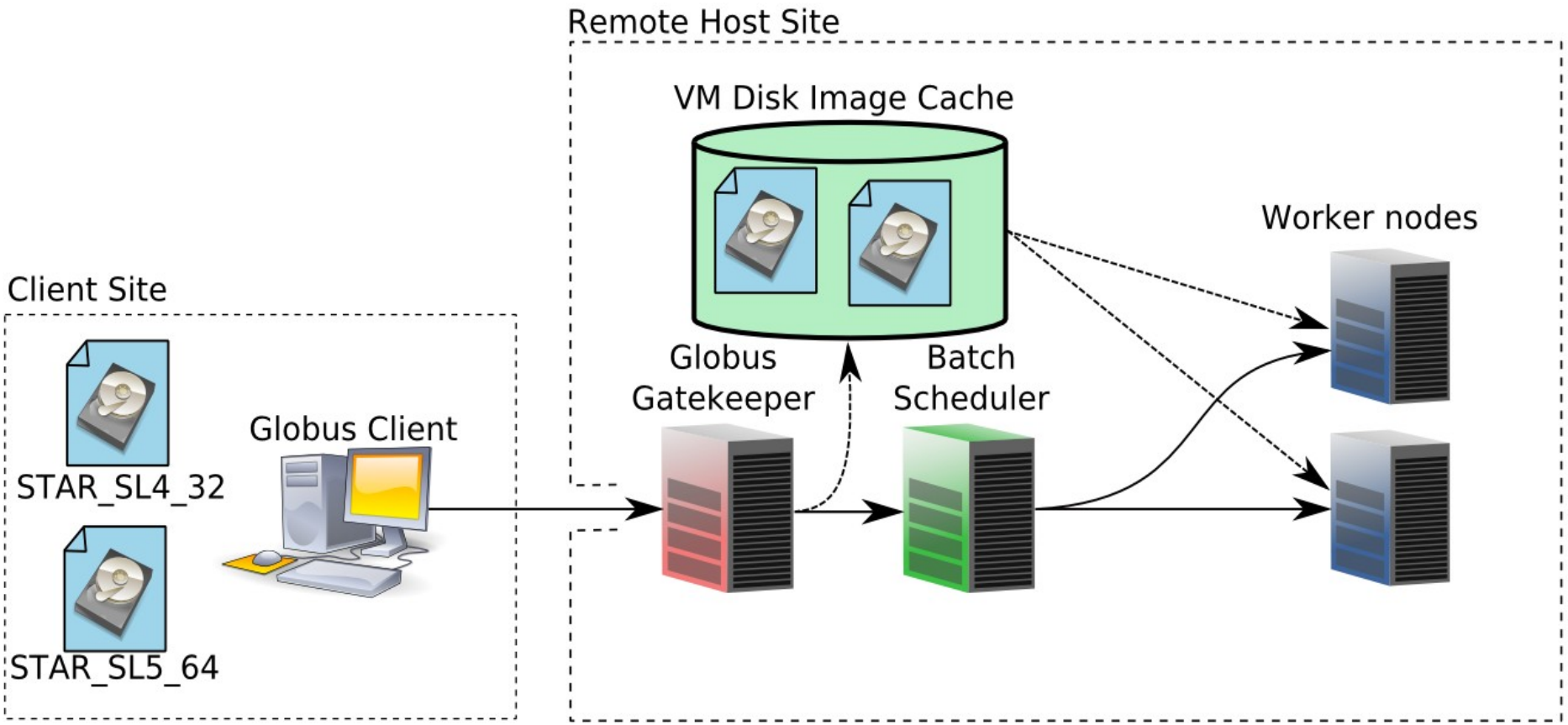
Host sites prepare site specific **base images** with different operating systems with contextualization pre-configured.



Problems with site specific base images:

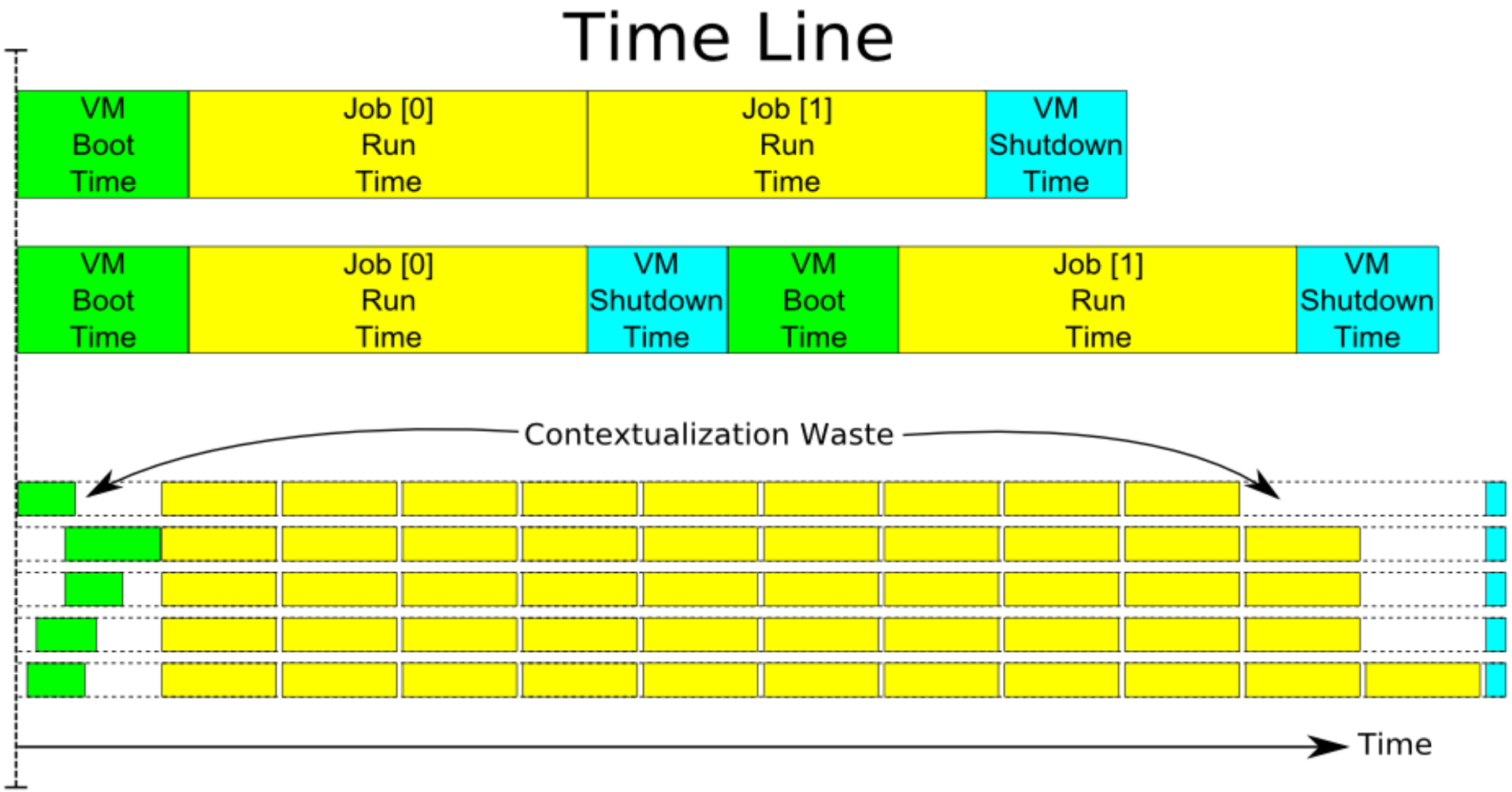
- Not being able to get a base image for the OS you want puts you **back to square one !**
- Host sites can not compose an infinite number of base images (usually very limited).

VM Image Management



Disk image files are usually a few GB, however all worker nodes generally are identical, so will only have to be uploaded at most once per request (group of jobs performing same work)). Selecting which request runs under what image and the caching of images should be the responsibility of a VM disk management system. So far the Globus Nimbus toolkit is the only package that we have encountered that performs this function.

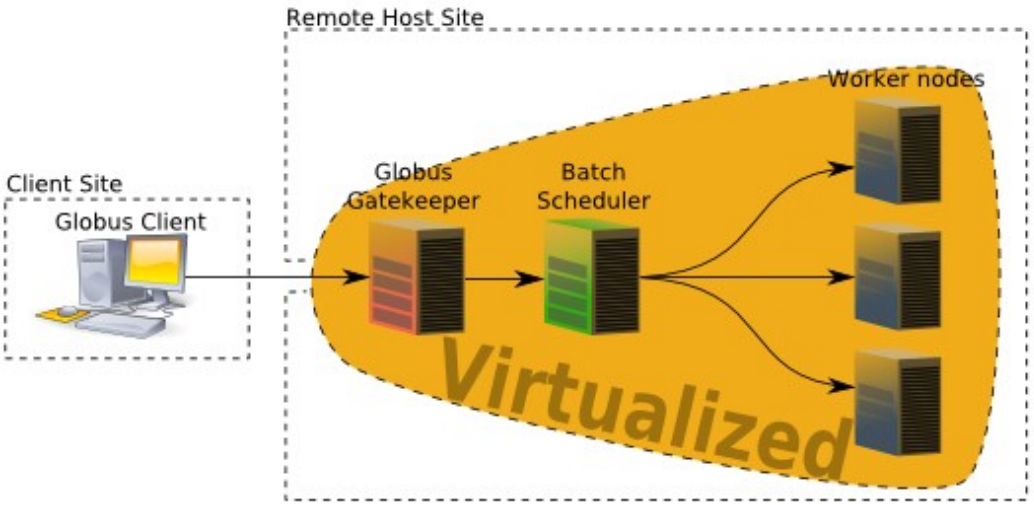
Efficiency of Different Running Models



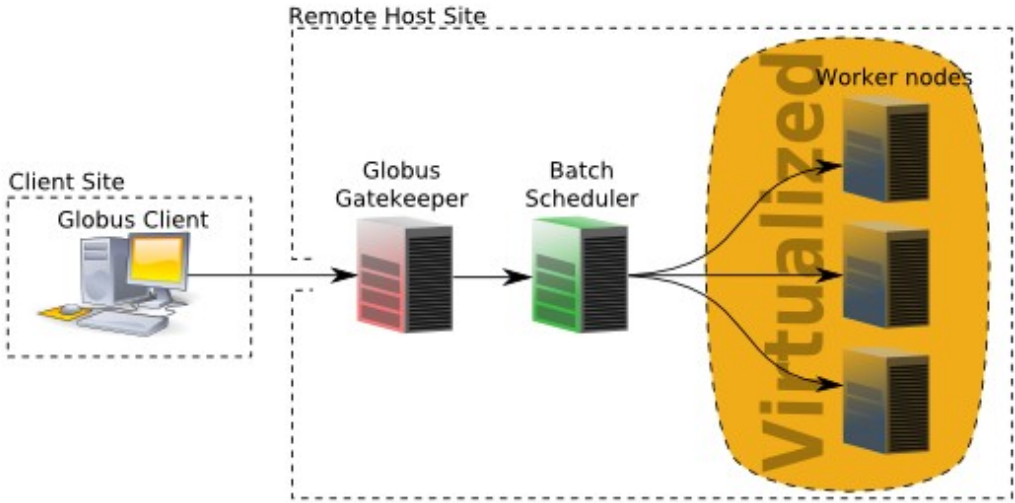
- On some models jobs can not start to run until the whole cluster is contextualized.
- Contextualization will make boot time longer depending on services started.

3 Models

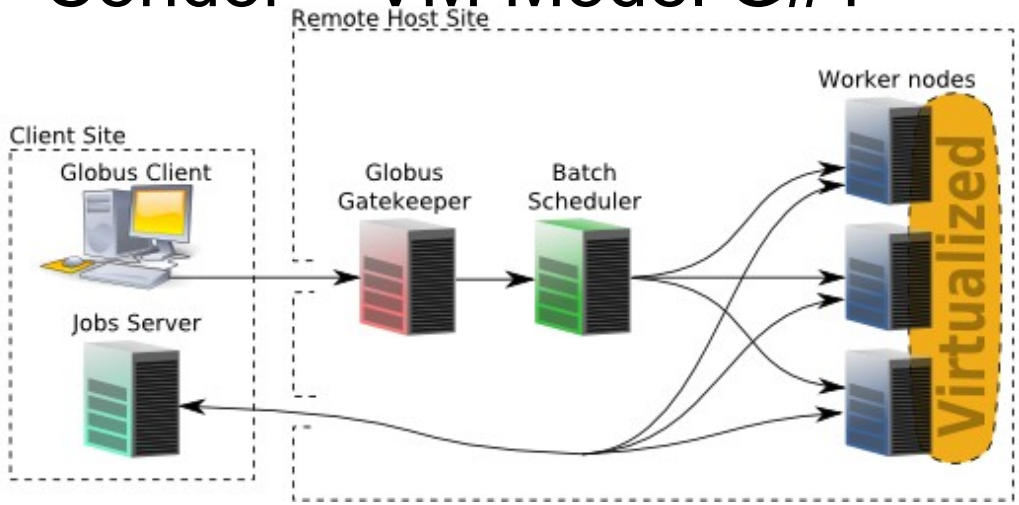
Amazon EC2 with Nimbus Interface



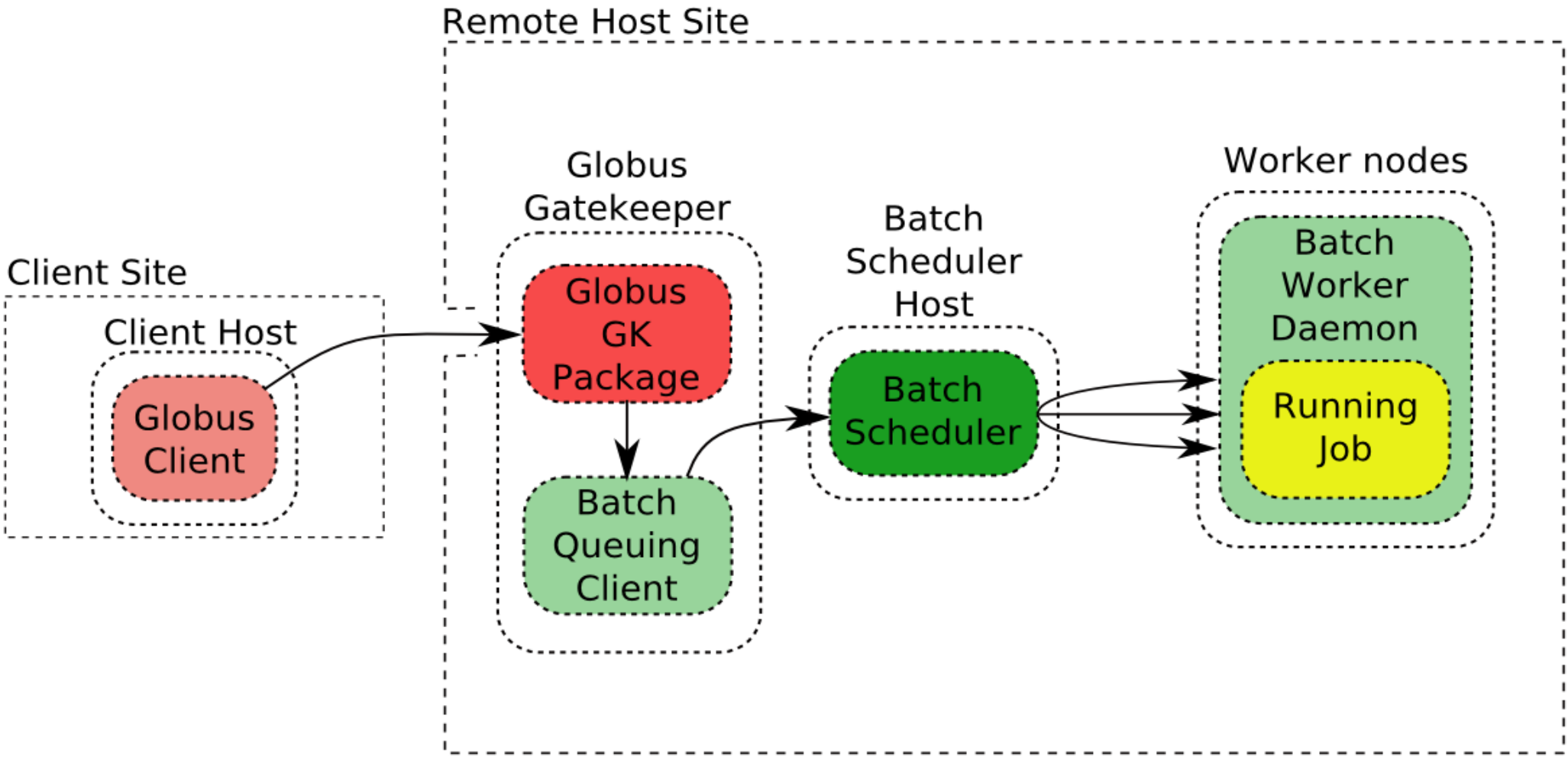
Clemson Model Cl#1



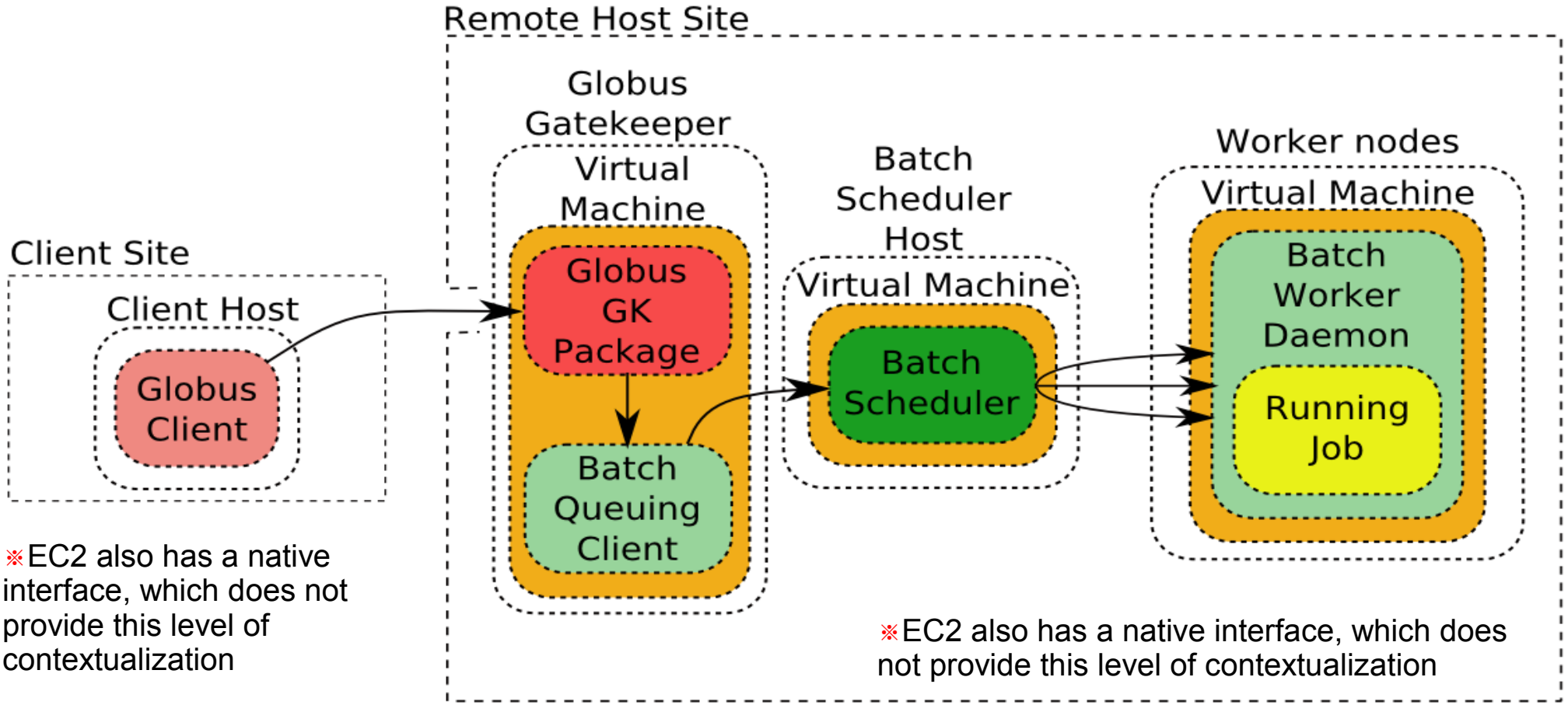
Condor – VM Model G#1



Non-Virtualized Grid Model (VDT/OSG)



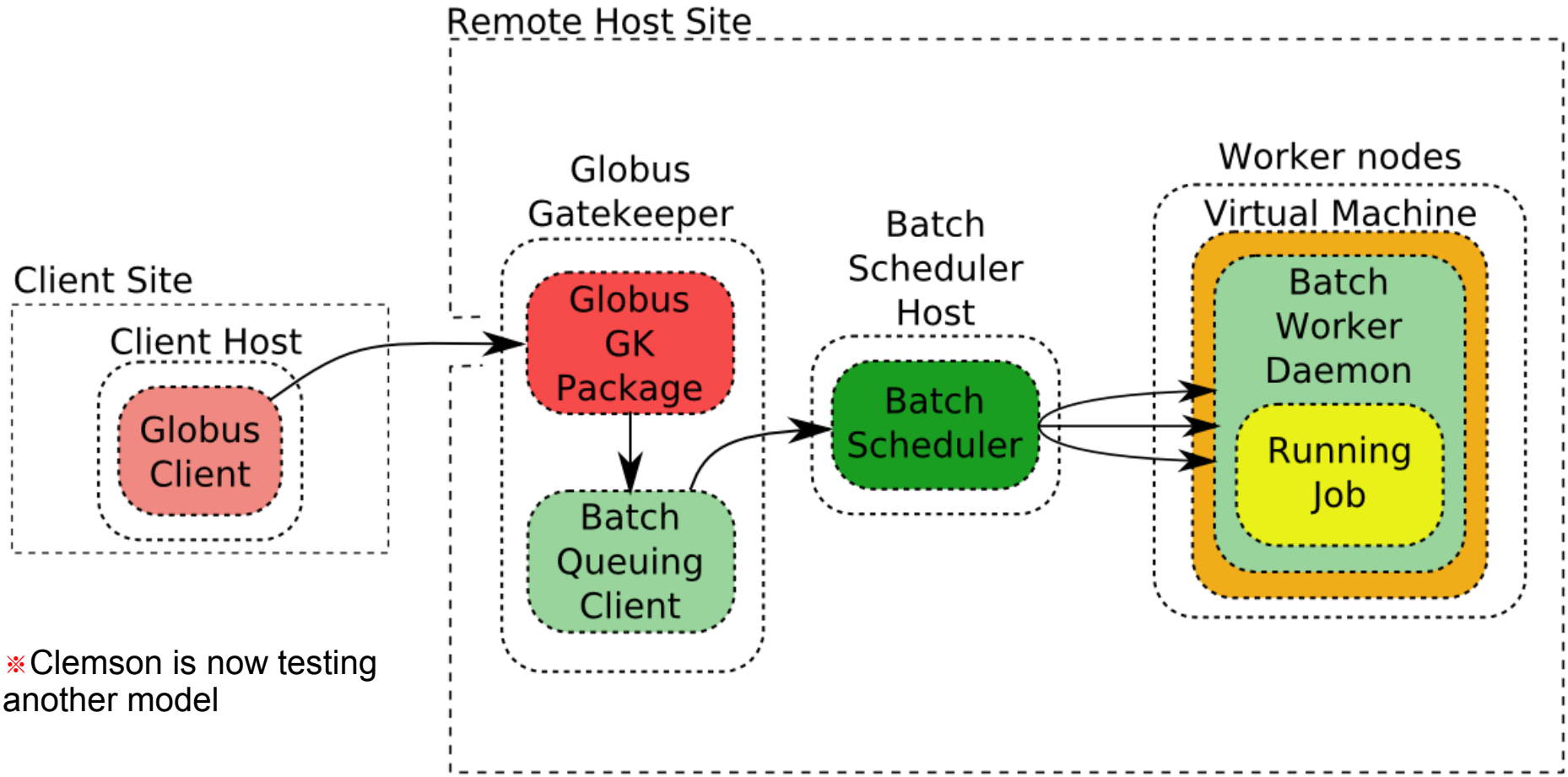
Amazon EC2 With Nimbus Interface Model



Pro	Con
-Guarantee on the number parallel slots (***not a hard requirement HENP (embarrassingly parallel)***) -Runs one job after the other without needing to boot up a new VM	-Base images need to be provided by host site -Contextualization waste on start-up and shutdown

◀-Submitting site is managing everything▶

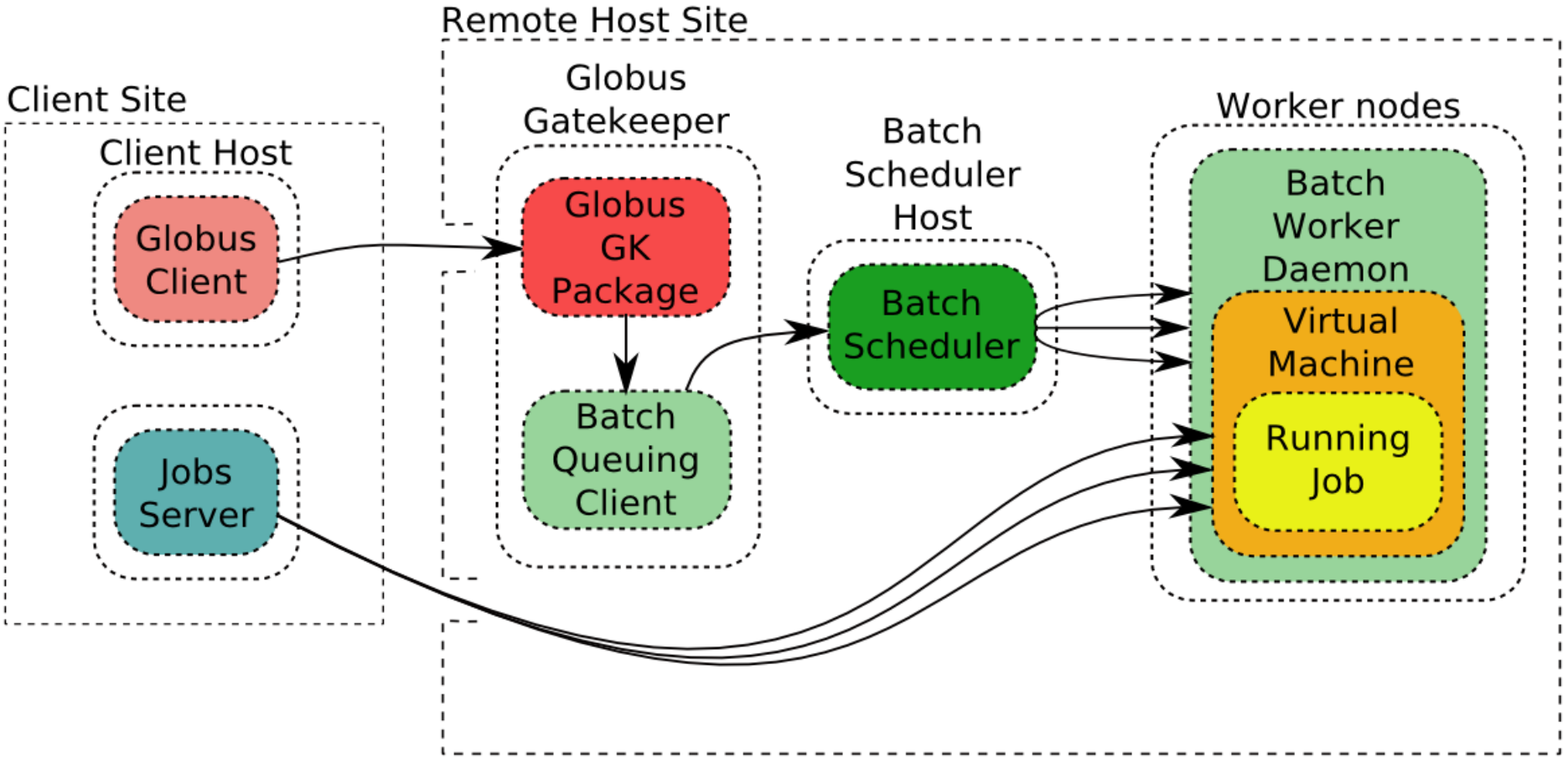
The Clemson Model CI#1



※Clemson is now testing another model

Pro	Con
-Most transparent to the user	-Batch worker MUST be supported by VM OS -Batch worker installed by host site into image (this is a lot of work for the host site)

Condor – VM Model G#1



Pro	Con
-Can run a large variety of images *(No site specific base image needed, no contextualization)*	-User must be trusted to shutdown the VM -User must figure out how to pull job in -Booting for each job is inefficient (multi-job submission framework must be supplied by user)

Conclusions

- Cloud Computing offers reproducibility
- Different models shift the responsibility of managing components between the submitters and host sites.
- The models offer trade-offs between portability and ease of use
- What would be the ideal model ?
  - Base Images and modifying user customization require significant effort from both host site and users. Testing each model is a significant effort.
  - Clemson model works best for end-users / VO:
    - Additions needed would be (wish list) :
      - Provide users a batch worker client they can easily install in a wide selection (Linux, Unix, Windows) of images (standardize).
      - Image management
      - Standardize submission interface across the grid
        - JLD to associate image with Job

End
Questions

Extraneous Slides

http://www.anl.gov/Media_Center/News/2009/news090402.html

Nimbus and cloud computing meet STAR production demands - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.anl.gov/Media_Center/News/2009/news090402.html

Google

Nimbus and cloud computing meet S...

Argonne NATIONAL LABORATORY

Newsroom

U.S. DEPARTMENT OF ENERGY

About Argonne | Research | Community/Environment | Newsroom | Working at Argonne | Search Argonne ...

Newsroom > News > 2009 >

- Pressroom
- Experts Guide
- News Releases
- Publications
- Internship

Nimbus and cloud computing meet STAR production demands

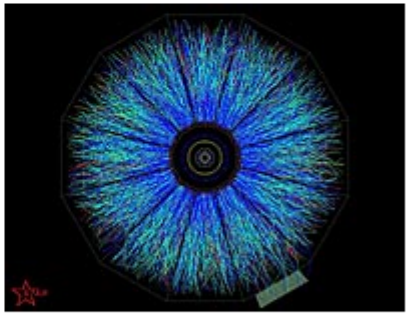
SHARE

(Download printer-friendly, PDF version.)

ARGONNE, Ill. (April 2, 2009) — The advantages of cloud computing were dramatically illustrated last week by researchers working on the STAR nuclear physics experiment at Brookhaven National Laboratory's Relativistic Heavy-Ion Collider. New simulation results were needed for presentation at the Quark Matter physics conference; but all the computational resources were either committed to other tasks or did not support the environment needed for STAR computations. Fortunately, working with technology developed by the Nimbus team at the U.S. Department of Energy's (DOE) Argonne National Laboratory, the STAR researchers were able to dynamically provision virtual clusters on commercial cloud computers and run the additional computations just in time.

Nimbus is an open source cloud computing infrastructure that provides tools allowing users to deploy virtual machines on resources, similar to Amazon's EC2, as well as

RESOURCES



A view of one of the first full-energy collisions between gold ions at Brookhaven National Laboratory's Relativistic Heavy Ion Collider, as captured by the Solenoidal Tracker At RHIC (STAR) detector. The tracks indicate the paths taken by

Find: Next Previous Highlight all Match case

Done

http://www.opensciencegrid.org/Clouds_make_way_for_STAR_to_shine

Clouds make way for STAR to shine - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.opensciencegrid.org/Clouds_make_way_for_STAR_to_shine

Open Science Grid

Search OSG using Google: search... GO

Search OSG at Work: search... GO

A national, distributed computing grid for data-intensive research.

Home

About the OSG

Learn About Us

What We're Doing

Getting Started with OSG

Contact Us

OSG at work

Calendar of Events

Collaborative Twiki and Chat

Technical Documentation

Security

Software

Education

Home > Clouds make way for STAR to shine

Clouds make way for STAR to shine

With local computers busy and the grid unable to guarantee workable resources on very short notice, the STAR experiment at Brookhaven National Laboratory turned to cloud computing to meet a tough deadline. Only two weeks before the biggest conference in the field, the team started a set of complex simulations and finished in time to present results.

"This is the very first time cloud computing has been used in our field for scientific production work with full confidence in the results" said Jérôme Lauret, who heads the STAR team. "This is a breakthrough."

Jérôme Lauret, left, and Kate Keahey, right. *Images courtesy of J.Lauret and ANL, respectively.*

STAR studies fundamental properties of nuclear matter as it exists in a high-density state called a Quark Gluon Plasma. This plasma decays into elementary particles, emitting narrow streams of fast-moving particles called jets. Scientists infer information from the jets about the conditions and properties of this plasma.

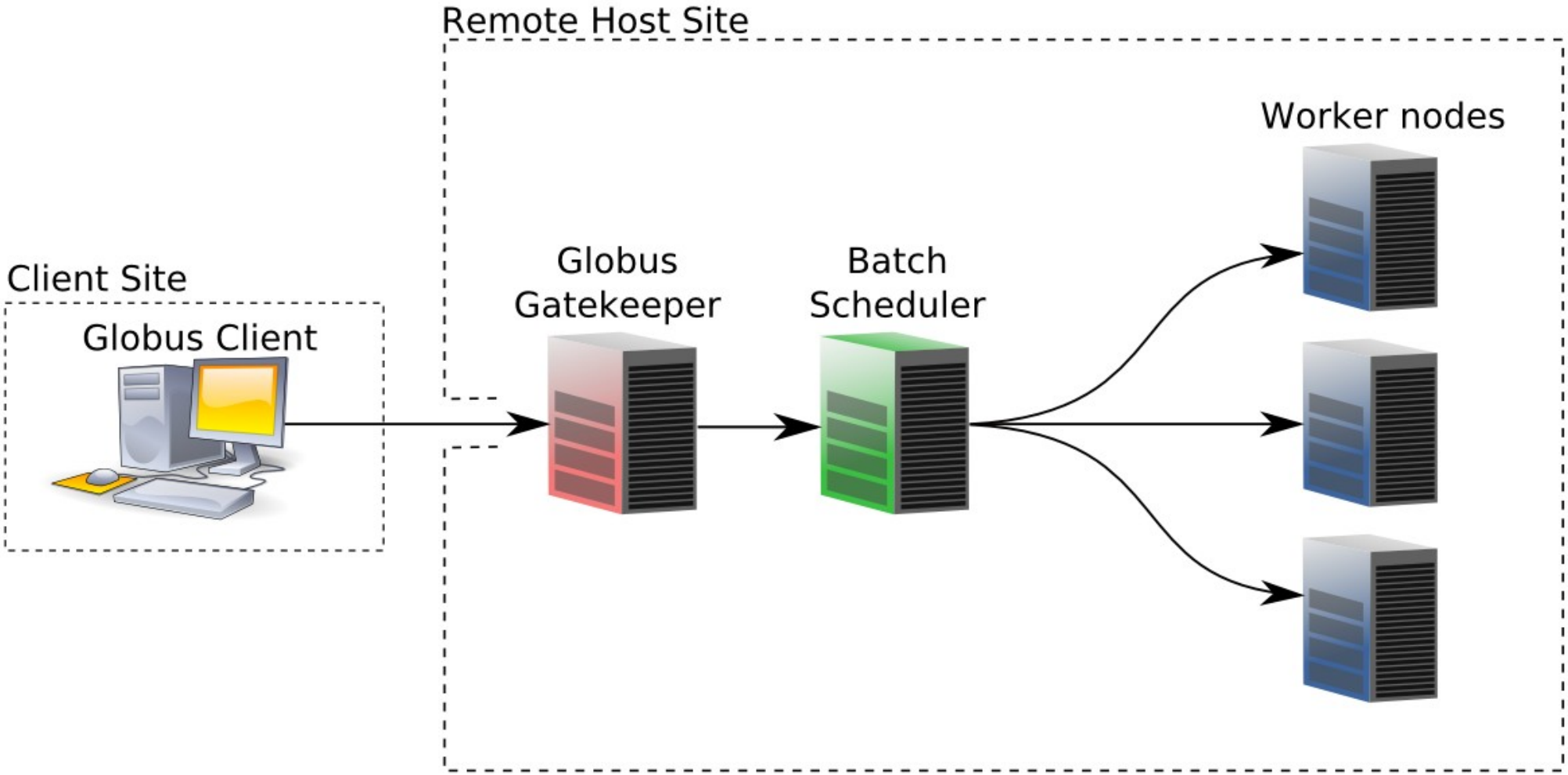
The STAR team had recently succeeded in reconstructing jets in the collisions of gold nuclei. Achieving and verifying this result required analysis of large data samples and simulations, and other compute-intensive tasks. STAR resources running the Open Science Grid software provided for much of this computation. However, behind the two-week, pre-conference scramble stands a collaboration begun a few years ago between STAR and the Nimbus team at Argonne National Laboratory, led by Kate Keahey.

Nimbus provides cloud computing tools that include the Nimbus Workspace Service, an open source

Find: Next Previous Highlight all Match case

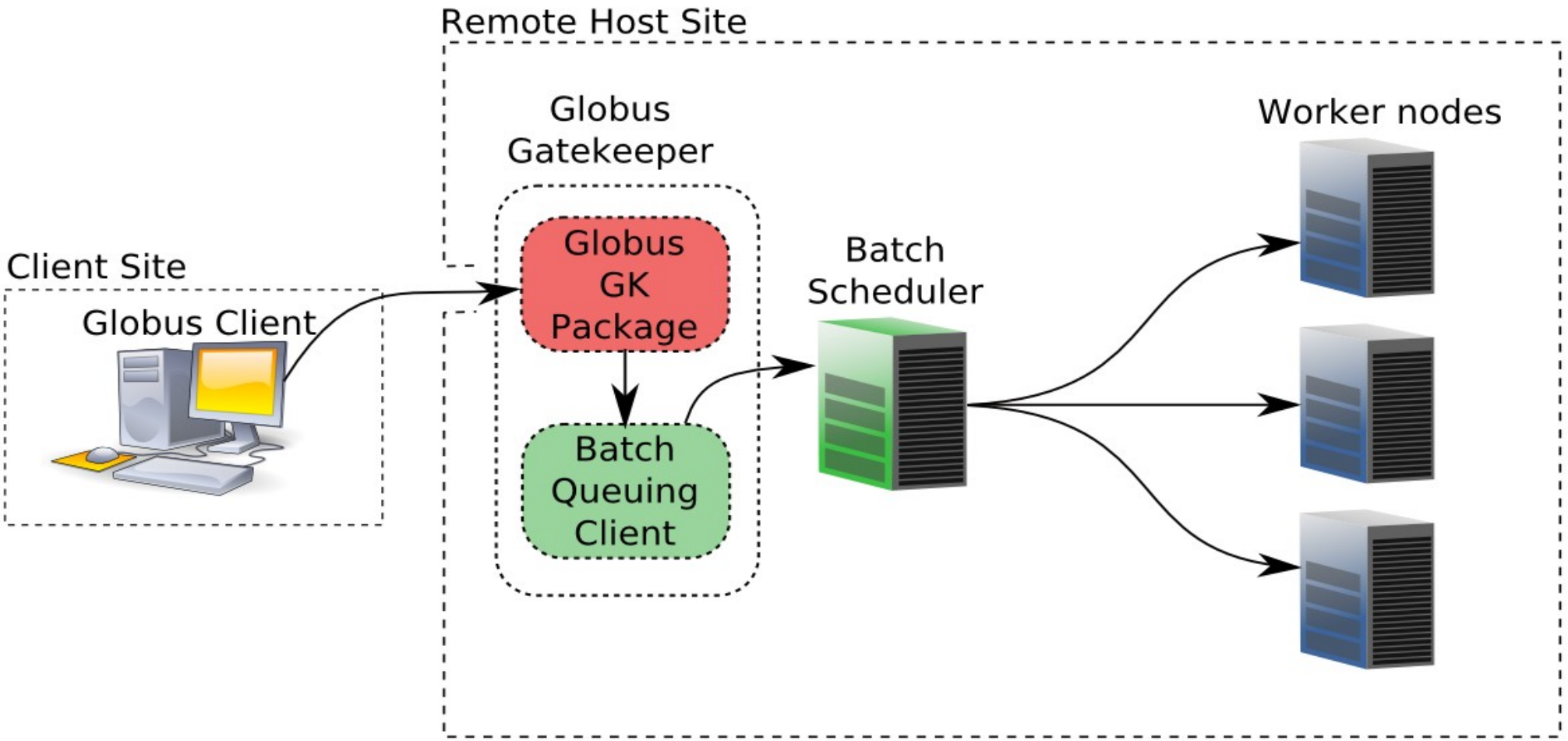
Done

Non-Virtualized VDT/OSG Model



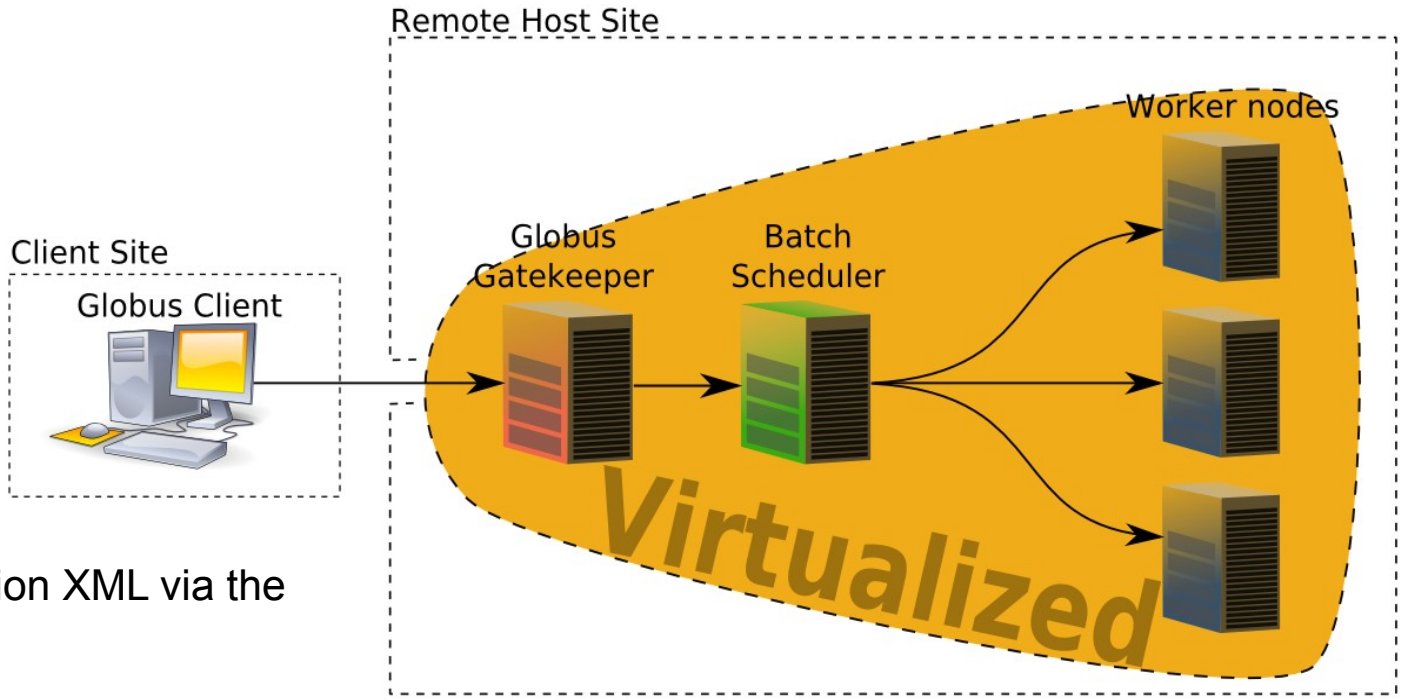
Nothing New Here

Taking a Look Inside (detail view)



EC2 with Nimbus Interface Model

Model: Whole Site is virtualized



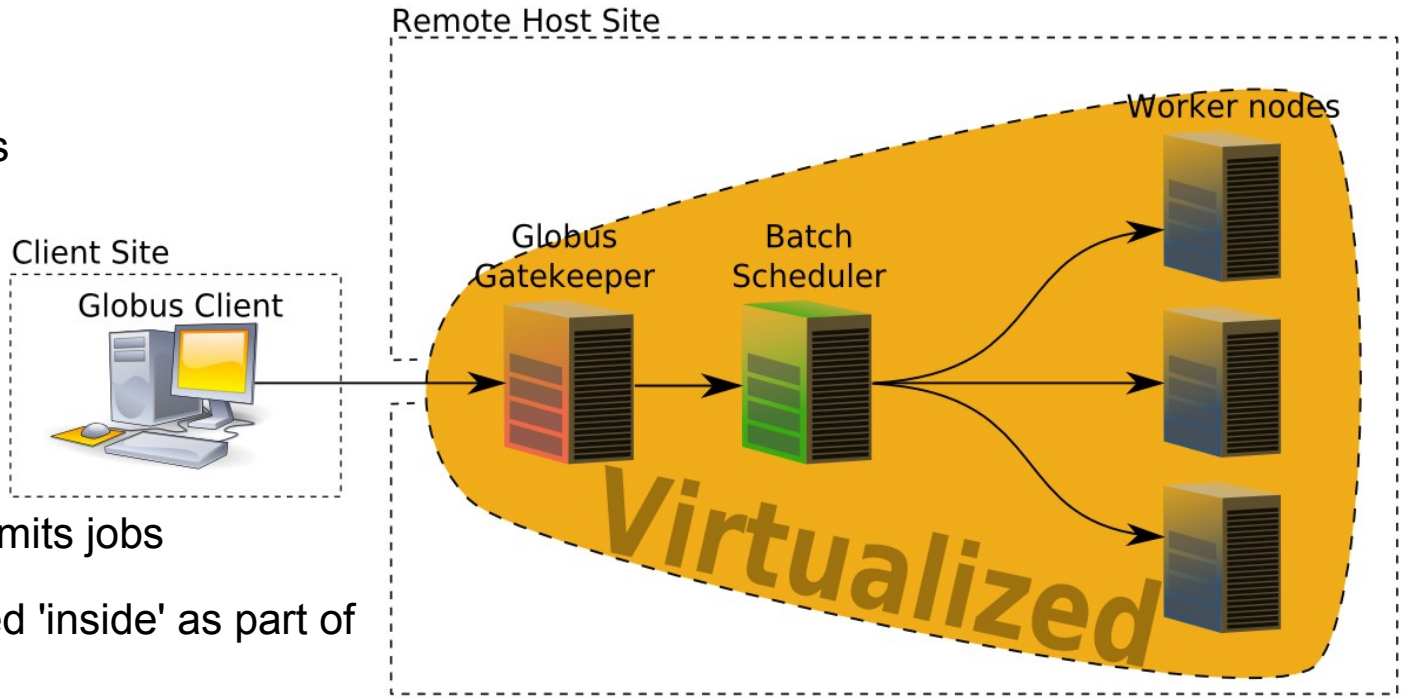
- User submits a cluster description XML via the Nimbus Client Toolkit
  - Includes pointers to GK image and worker node image, and the number of worker nodes to contextualize
- After contextualization user submits jobs
  - batch system and GK was deployed 'inside' as part of a contextualization
- When finished cluster is shut down via the Nimbus Client

※EC2 also has a native interface, which does not provide this level of contextualization

EC2 with Nimbus Interface Model

Model: Whole Site is virtualized

- User submits a cluster description XML via the Nimbus Client Toolkit
  - Includes pointers to GK image and worker node image, and the number of worker nodes to contextualize

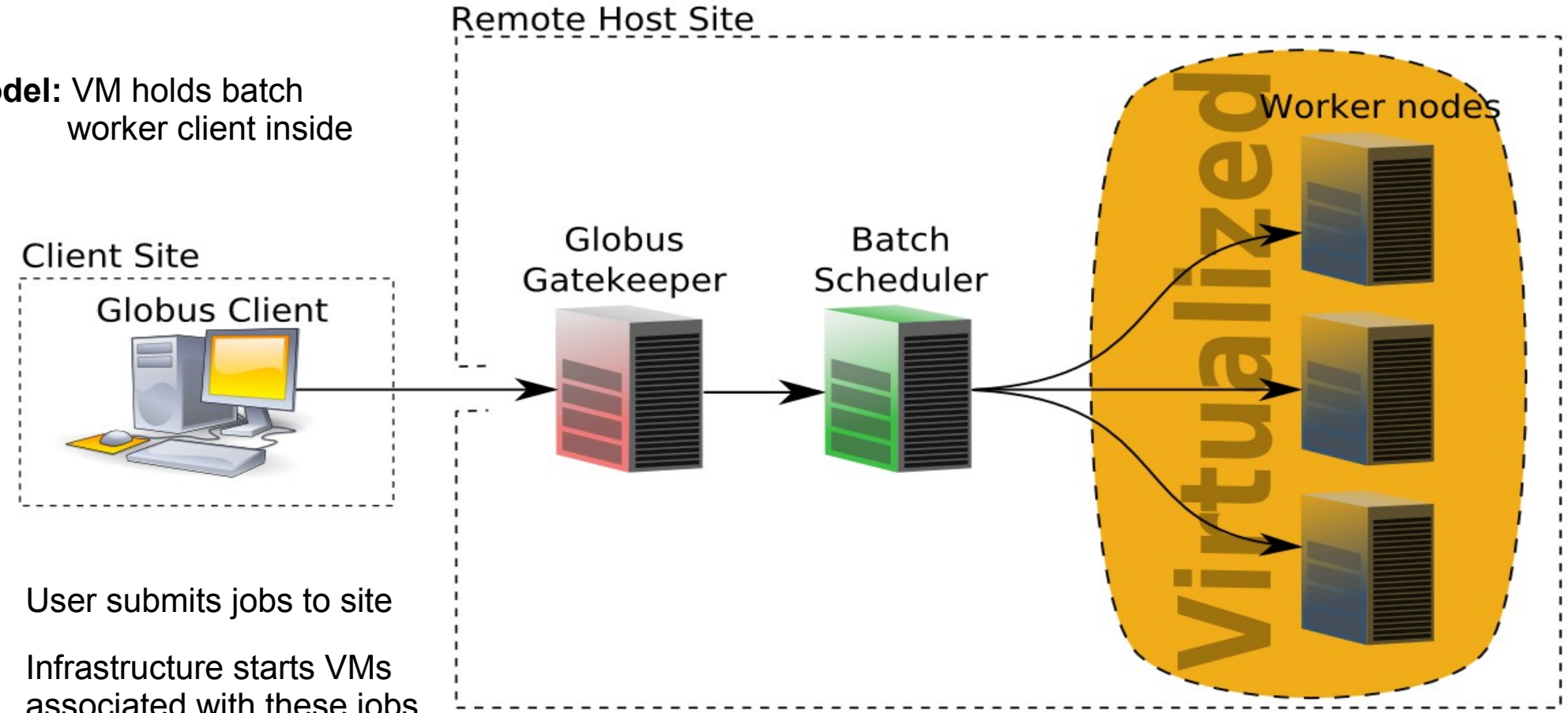


- After contextualization user submits jobs
  - batch system was deployed 'inside' as part of a contextualization
  - we start WN and a head node with pre-package Grid stack for convenience (STAR/Nimbus specific implementation)
- When finished cluster is shut down via the Nimbus Client
  - *cannot shutdown until the last jobs finishes*

※EC2 also has a native interface, which does not provide this level of contextualization

The Clemson Model CI#1

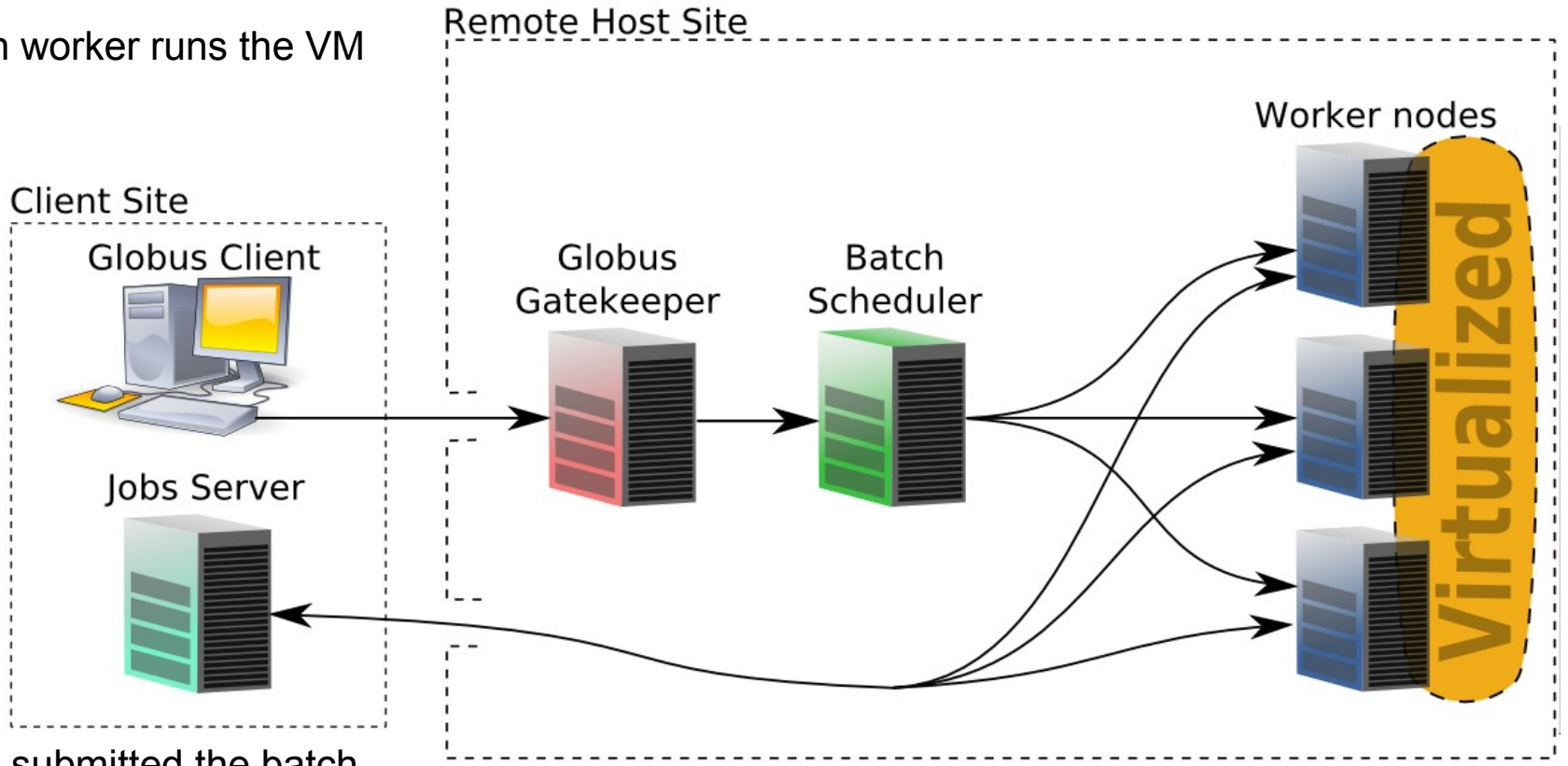
Model: VM holds batch worker client inside



- User submits jobs to site
- Infrastructure starts VMs associated with these jobs
- Batch worker client inside
  VM registers itself with batch scheduler as worker meeting the resource requirements of the jobs.
- Jobs are processed.
- When no more jobs with these requirements are queued, the infrastructure shuts down the VM

Condor – VM Model G#1

Model: The batch worker runs the VM



- For each job submitted the batch worker starts a VM
- The VM must have “some way” of pulling in a job or the job must already be installed inside the VM
- When finished the job must shut down the VM

※If One VM could run multiple “jobs” via a pilot and remote queue however the submitters software must support this.

※Condor is now testing a publish / subscribe model.

Conclusions Summary

	Nimbus / EC2	Clemson	Condor-VM / GLOW
Contextualization scope	whole cluster	node	(none) one job
Contextualization needed	heavy	light	Very light
Base Images(site specific)	needed	limited need	not needed
Batch system managed by:	submitter	host site	host site
Batch worker managed by:	submitter	submitter	host site (none inside VM)
GK managed by:	submitter	host site	host site
Has image management	yes	no	no
VM associated with:	cluster	user	job
Thanks To:	Kate Keahey & Tim Freeman Argonne National Laboratory University of Chicago	Michael Fenn Sebastien Goasquen Clemson University	Miron Livny Greg Thain Jan Balewski (testers) Matthew Walker (testers) University of Wisconsin–Madison